Cognition 127 (2013) 99–104

Contents lists available at SciVerse ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Brief article

Learning different light prior distributions for different contexts

Iona S. Kerrigan, Wendy J. Adams *

Psychology, University of Southampton, Southampton SO17 1BJ, UK

ARTICLE INFO

Article history:
Received 13 June 2012
Revised 26 November 2012
Accepted 17 December 2012
Available online 30 January 2013

Keywords:
Shape-from-shading
Human visual perception
Light priors
Context specific learning

ABSTRACT

The pattern of shading across an image can provide a rich sense of object shape. Our ability to use shading information is remarkable given the infinite possible combinations of illumination, shape and reflectance that could have produced any given image. Illumination can change dramatically across environments (e.g. indoor vs. outdoor) and times of day (e.g. mid-day vs. sunset). Here we show that people can learn to associate particular illumination conditions with particular contexts, to aid shape-from-shading. Following a few hours of visual–haptic training, observers modified their shape estimates according to the illumination expected in the prevailing context. Our observers learned that red lighting was roughly overhead (consistent with their previous assumption of lighting direction), whereas green lighting was shifted by 10°. Greater learning occurred when training for the two contexts (red or green light) was intermingled rather than when it was sequentially blocked.

1. Introduction

Humans cope with reddish illumination at sunset or flickering coloured lights at the disco – managing to decompose shading patterns into reflectance and shape variations – but how? Our impressively robust ability to estimate our surroundings, given complex and ambiguous retinal input relies heavily on prior knowledge – we bias perceptual estimates toward the most likely scenes. For example, we bias estimates of illumination direction toward overhead (e.g. Adams, 2007; Kleffner & Ramachandran, 1992) and estimates of surface shape toward convexity (Adams & Mamassian, 2004; Langer & Bülthoff, 2001) in alignment with the statistics of our environment (Potetz & Lee, 2003). Such assumptions, or 'priors', facilitate the notoriously underconstrained problem of recovering shape-from-shading. Here we investigate whether observers can further refine this process by learning that particular illumination conditions are more likely in particular contexts.

For optimal performance, humans should (i) respond to long-term changes in scene statistics by updating their priors and (ii) select the correct prior for a given context. We know that humans do the former: in contrast with chickens (Hershberger, 1970), human observers change their light prior in response to appropriate haptic (Adams, Graf, & Ernst, 2004) or visual feedback (Adams, Kerrigan, & Graf, 2010). Here we ask whether humans also do the latter: can we learn different prior assumptions for different contexts? There is no clear consensus: although Adams et al. (2004) found that a modified light-prior generalised to novel stimuli, Adams et al. (2010) noted that modified light-priors were retained for several weeks beyond training, after observers had returned to their normal environment, in which lighting was presumably, on average, overhead. This latter finding suggests that observers learnt separate, context-dependent light priors, with the experimental set-up acting as a contextual cue.

Here we ask whether humans can learn two light priors, each invoked by a different illumination colour. To induce colour-dependent learning, visual–haptic feedback was modulated by the simulated illumination colour: when scenes were illuminated by red light, feedback was consistent with the observer's baseline light prior distribution. In contrast, under green illumination, feedback was consistent with a new lighting distribution.

* Corresponding author. Tel.: +44 02380 593629.
E-mail addresses: i.s.kerrigan@soton.ac.uk (I.S. Kerrigan), w.adams@soton.ac.uk (W.J. Adams).

0010-0277
http://dx.doi.org/10.1016/j.cognition.2012.12.011

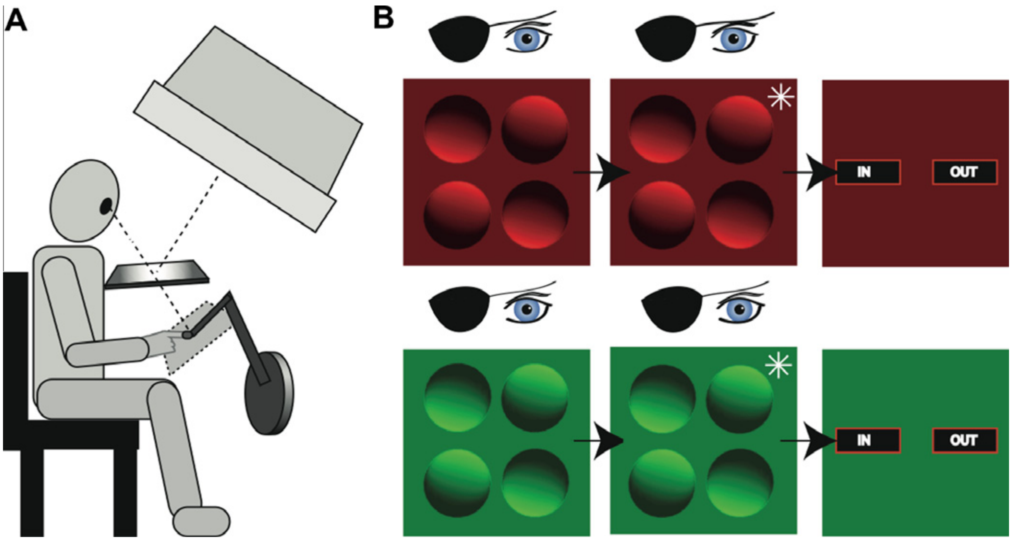

Fig. 1. Apparatus and visual test trials: (A) The visual–haptic experimental set-up. (B) Examples of visual-only test trials: the simulated lighting is either red (upper row) or green (lower row). Observers briefly viewed the four shaded discs (total presentation time 1.2 s, target object cued after 600 ms) before indicating whether the cued object was concave or convex (in or out). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2. Materials and methods

2.1. Apparatus and stimuli

Observers simultaneously viewed and felt virtual objects (see Fig. 1a). Haptic scenes were presented via a 'thimble gimbal' attached to a force-feedback device (Ghost libraries, PHANToM, SensAble Technologies). Visual stimuli (Figs. 1b and 2d–g), generated using OpenGL, were presented via a front-silvered mirror. Their perceived location (at a visual distance of 56 cm) matched the location of the haptic stimuli, giving the impression of a single visual–haptic scene. A headrest and bite bar maintained head position and an eye patch eliminated binocular depth cues. The room was completely dark, other than the light emitted by the visual display.

2.2. Visual test trials

Pre- and post-training trials contained solely visual (no haptic) information. Observers viewed four shaded discs, each subtending 5.6° and offset from the screen's centre by 5.3° (see Fig. 1). Each disc was consistent with a hemisphere squashed in depth by a factor of 2, illuminated by a distant light source. The slant of the light source (the angle between the lighting vector and the screen normal) was 68.2°. The light source tilt (the angle between the projected lighting vector and the vertical axis in the plane of the screen, θ) varied across trials. This illumination tilt, with object shape (convex vs. concave) determined the shading orientation of each disc. Within each trial, one, two or three discs had a shading gradient direction of θ and the remaining disc(s) had a shading gradient of $\theta + 180°$, such that observers generally perceived both convex and concave objects to be present. The simulated scene was white, with either a red or green simulated light source although stimuli were equally consistent with red and green scenes illuminated by white light.

Observers judged the shape (concave vs. convex) of one object (cued by a star). The observer's light prior was estimated from the set of 288 visual trials (24 equally spaced θ values $\times$ 2 colours $\times$ 6 repetitions), lasting approximately 10–15 min (see Fig. 2a).

2.3. Training trials

Visual–haptic training was similar to that used previously (e.g. Adams et al., 2004, see Fig. 2d–g). Observers viewed four shaded discs (as in test trials), but also explored the scene haptically by running a finger (in a thimble gimbal) over the simulated objects. This haptic information disambiguated each object's shape, and thus also the lighting direction. However, the relationship between shading orientation and haptic shape depended on colour (see Fig. 2b). On 'red' trials, stimuli were consistent with the observer's baseline light prior; haptic shape matched the observer's pre-training shape responses. On 'green' trials, however, the lighting direction was drawn from a range shifted by ±30° relative to the observer's baseline prior (13 observers were assigned a +30° shift, 13 a −30° shift). Thus, on 'green' trials, some objects previously perceived as convex now felt concave, and vice versa.

It is important to note that haptic feedback did not introduce an association between colour and shape; $p(\text{haptically convex}|\text{green}) = p(\text{haptically convex}|\text{red})$. Rather, for perception to become aligned with haptic feedback, the observer would have to learn a relationship between illumination direction and colour.

After haptically exploring the scene for a minimum of 7 s, including 'touching' all four objects, the observer pressed a button to continue. One of the objects then appeared visually (without haptics) in the centre of the

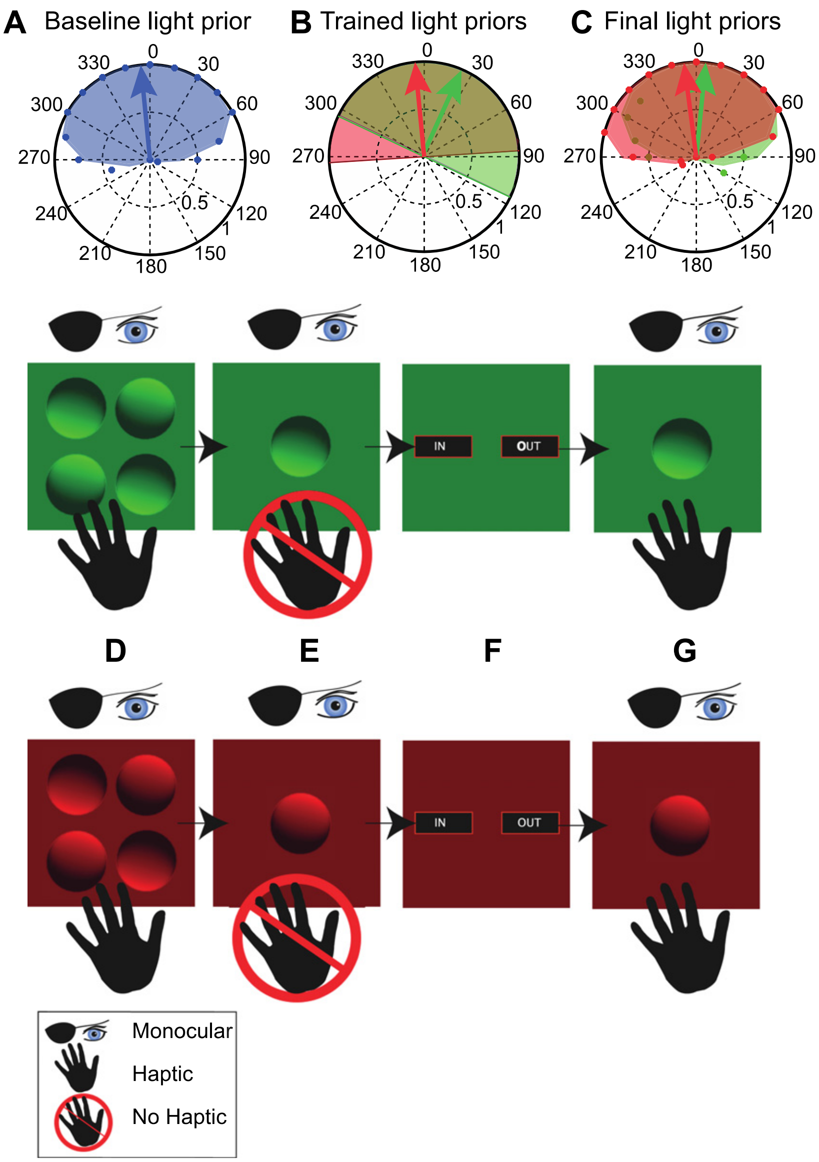

Fig. 2. Visuo-haptic training trials: (A) Proportion of responses that were convex, as a function of shading orientation, for one observer. The shaded region represents the fitted baseline light prior distribution, with the mean of the light prior given by the blue arrow. (B) Two trained light priors. The red region indicates the range of stimulus shading orientations that 'felt' convex during red visual–haptic training trials – it matches the range of orientations perceived as convex at baseline. The green region indicates the orientations that 'felt' convex during green visual–haptic trials. (C) Proportion of convex responses on red and green final test trials. The red and green shaded regions represent the fitted light priors for each colour with the mean of each light prior given by the red and green arrows respectively. (D–G) Schematic representation of a training trial: (D) Observers explored the scene both haptically and visually and then (E) viewed a single disc for 1 s before (F) judging its shape. (G) Viewing and touching the single stimulus provided feedback. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

screen for 1s and the observer judged its shape (convex/concave). By subsequently viewing and touching the object, observers gained feedback on their response. Each training set comprised 224 visual–haptic training trials (48 equally spaced θ values $\times$ 2 colours $\times$ 2 repetitions + 2 extra repetitions of 8 θ values within conflict regions $\times$ 2 colours), lasting approximately 60–90 min.

There is some evidence that people and animals learn to discriminate between two contingencies more quickly when trials are intermixed than when they are blocked (e.g. Honey, Bateson, & Horn, 1994; Mitchell, Nash, & Hall, 2008). To identify whether a similar advantage is observable for our context-dependent learning task, we assigned observers to either an (i) intermingled or (ii) blocked variant. In the first variant, red and green trials were randomly intermingled throughout test and training. In the second, colour was fixed within blocks of 24 trials.

2.4. Procedure

On day 1, each observer completed a set of visual-only trials, followed (after a short break) by a train-test session (one set of visual–haptic training trials, one set of visual test trials). On day 2, they completed two train-test sessions, separated by at least 1 h.

2.5. Participants

Twenty-six naïve observers completed the experiment (intermingled variant: 10 participants; blocked variant: 16 participants). All had normal or corrected-to-normal visual acuity and normal colour vision. Participants gave informed written consent and the local ethics committee approved the study.

2.6. Possible outcomes

What might our observers learn from the visual–haptic training? What colour-dependent or colour-independent changes in shape perception might be seen? First, our observers might show no learning: if colour were ignored as a 'nuisance' variable, haptic feedback would appear noisy and inconsistent and thus might be discounted. Second, observers might ignore colour, but still modify their behaviour to reflect the aggregate of all feedback. Their light priors would thus move toward the average of the two trained lighting distributions, irrespective of stimulus colour. Finally, observers may learn (consciously or unconsciously) that particular colours are associated with particular lighting distributions. This would allow them to apply different prior distributions over lighting direction in different colour contexts. This context-specific learning would result, post-training, in different measured light priors for different coloured test stimuli – the same shading orientation would induce different perceived shapes under different illumination colours.

3. Results

Light priors were estimated by fitting a simple Bayesian model to each observer's test and training data (see Adams et al., 2010; essentially, the peak of the light prior is given by the peak of the 'convex' responses). We first checked whether the perceived shape of red and green stimuli differed prior to training. Three observers were excluded (one from the intermixed condition, two from the blocked condition) as their red and green baseline priors differed (*ps* = .015; .015; and .003, from bootstrapping). A single baseline light prior was estimated for each remaining observer, using their combined red and green pre-training data (M = −9.61°, SD = 13.61°, across observers). All subsequent data were separated by colour to estimate colour-specific light priors.

Results are shown in Fig. 3. Training had a significant effect on shape perception ($F(3, 63) = 6.61$, $p = .003$, $\varepsilon = .67$, G–G correction for non-sphericity, from 3 factor ANOVA (amount of training, training type and illumination colour), partial $\eta^2 = .24$). As training progressed, observers' light prior distributions moved toward the trained lighting direction. Significant learning occurred after two sets of training (normalised baseline light prior = 0°, vs. mean penultimate and final light-priors $\mu = 7.71°$ and $\mu = 9.45°$, $p = .012$ and $p = .042$ respectively, from Bonferroni corrected comparisons).

Importantly, our observers did show some context dependent learning: light priors were shifted significantly further from baseline when measured with green than with red test stimuli ($F(1, 21) = 5.28$, $p = .032$, partial $\eta^2 = 0.20$). Further analysis showed that this difference (green light prior – red light prior) reached significance at the first and third post-training tests (test 1: $t(22) = 2.26$, $p = .034$, $\mu_{\text{green}} = 4.38°$, $\mu_{\text{red}} = 2.20°$; test 3: $t(22) = 2.25$, $p = .035$, $\mu_{\text{green}} = 9.77°$, $\mu_{\text{red}} = 6.61°$). However, learning was not entirely context dependent: significant changes in light prior were observed for both illumination contexts by the end of training (green: $\mu = 9.77°$, red:

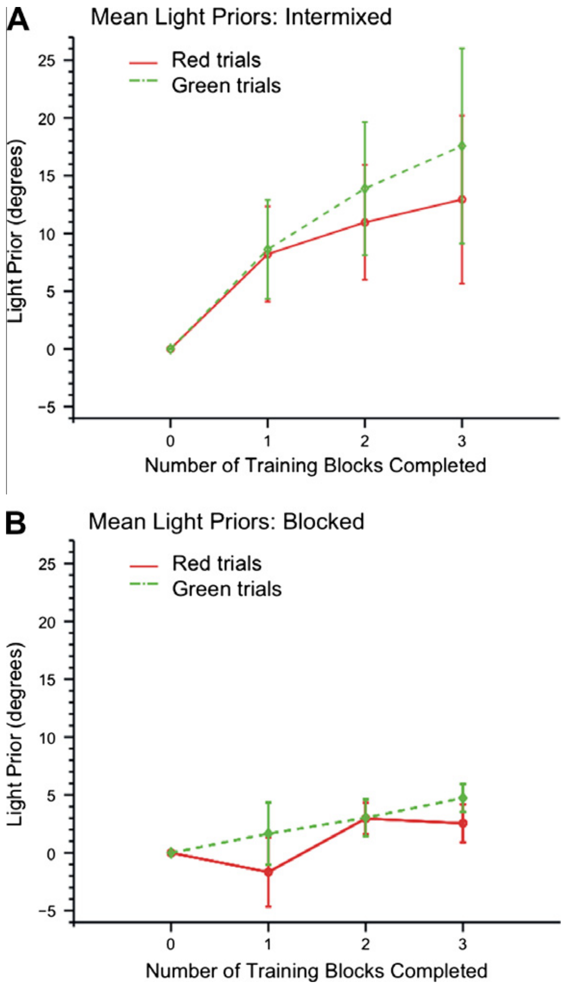

Fig. 3. Light priors before and after training: (A) in the intermixed condition and (B) blocked condition. To allow meaningful comparisons across observers, each observer's data were normalised by his or her baseline light prior, and light-priors for observers who trained with a −30° shift were multiplied by −1.

$\mu = 6.61°$, t-tests against zero: $p = .011$ and $p = .045$ respectively).

Participants trained via the intermixed condition showed significantly more learning than those trained via the blocked condition (15.25° vs. 3.64° in the final test session, significant main effect of training type (blocked vs. intermixed): $F(1, 21) = 4.31$, $p = .05$, partial $\eta^2 = 0.30$). This parallels findings from perceptual learning studies in which intermixed training results in better discrimination between stimuli than when training is blocked (e.g. Honey et al., 1994; Mitchell et al., 2008). The current study demonstrates that a similar advantage is seen when acquiring context-specific priors, perhaps because intermixing contexts avoids habituation of the contextual cue. No interactions were significant.

4. Discussion

We show that the visual system is able to learn and implement separate light priors for different contexts. After training, the perceived shape of ambiguous shaded objects was modulated by illumination colour. Our observers were able to switch between two different, colour-contingent light priors on a trial-by-trial basis. We are confident that our training affected visual perception, rather than creating a response bias: each test stimulus produced one of two, qualitatively different, shape interpretations (either convex or concave) and the observer simply reported which was perceived. From the observer's perspective our task was trivially easy: he/she simply reported an unambiguous, dichotomous percept. This can be contrasted with tasks involving a continuous response variable, where the potential for response bias can be problematic (see Haijang, Saunders, Stone, & Backus, 2006; Backus, 2009for further discussion). Furthermore, this colour-contingent perception was implemented unconsciously – at debrief, observers were asked to identify any differences that they had noticed between red and green stimuli, beyond the obvious difference in colour. Although some spurious differences were suggested, e.g. there were more trials of one colour than the other, observers did not identify any difference related to illumination direction, shading gradient or shape. This implicit learning is consistent with other evidence that observers can learn cue relationships of which they are unaware (Di Luca, Ernst, & Backus, 2010).

Did our observers learn a relationship between (i) illumination direction and illuminant colour, or between (ii) illumination direction and object colour? Our stimuli were equally consistent with white scenes under red/green illumination, or red/green scenes under white illumination. We suggest that the former is more likely; one can imagine illuminants at different locations providing differently coloured light. For example, the setting sun produces a more red/orange illumination than a midday sun, or a room may include red and green point light sources at different locations. In contrast, the latter interpretation requires that certain illumination directions are more likely when an object of a particular colour is present. In either case, however, separate priors over illumination direction have been learnt, one for each colour (of object, or illuminant).

Context-dependent use of priors has clear benefits: implementing a light prior that accurately reflects the lighting statistics of the current context will lead to more accurate shape judgements. Interestingly, however, the observed learning was not entirely context-specific; observers did not fully differentiate on the basis of colour but instead showed a combination of colour-specific and colour-independent learning. This may reflect robustness to temporary, perhaps spurious changes in cue contingencies; previous experience suggests that illumination direction and colour are not strongly correlated. In other words, this learning response may reflect not a limitation in our ability to learn, but an optimal strategy given the likelihood of such changes. This robustness has been modelled using Kalman filters where weight is given to both historical and current input (e.g. Burge, Ernst, & Banks, 2008). A similar logic applies to the incomplete learning found here (final average 'green' light-prior 9.77°, vs. trained light-prior of 30°) and in previous studies with comparable training, where observers learned only a third (11°) of the trained light prior shift (Adams et al., 2004).

Other research has investigated the priors over aspect ratio that contribute to orientation estimation (an elliptical retinal image may be perceived as a slanted circle). Knill (2007) demonstrated that observers use visual–haptic feedback to modify their prior on the aspect ratio of ellipses. Similarly to Adams et al. (2010), Knill noted that observers' priors did not readapt on exposure to the normal environment, suggesting that learning was specific to the laboratory context. Later, Seydell, Knill, and Trommershäuser (2010) demonstrated that observers can learn separate aspect ratio priors for different shapes (diamonds vs. ellipses) but appear unable to learn aspect ratio priors conditioned on object colour. The authors suggest that colour cannot be used to modulate observers' shape priors because colour is deemed to be unrelated to aspect ratio by the visual system: there is no ecological reason to link colour with aspect ratio. Our study shows that colour can act as a contextual cue; it may be that a relationship between illumination colour and illumination direction is deemed more plausible by the visual system.

A similar argument is presented by Michel and Jacobs (2007). They suggest that learning will be relatively easy when an existing relationship is modified (parameter learning). In contrast, learning a new relationship (structural learning) will be difficult or impossible: they tested whether observers could learn an association between illumination direction and stimulus depth, but concluded that they could not. Under this framework, our observers learned the distinction between colour contexts because colour and illumination direction are sometimes related in the real world: the relationship has ecological validity. However, it remains unclear whether, given enough training, observers would be able to learn two different light priors for contextual cues that are ecologically unrelated to illumination (e.g. object shape or texture). Ernst (2007) demonstrated that people can learn to associate low-level cues (luminance and stiffness) that have previously been unrelated. This suggests ecological validity is not a necessary condition for learning.

In broad agreement with Michel and Jacobs (2007), Backus and colleagues have shown that some cue associations are easier to learn than others (for a review see Backus, 2011). They asked which novel cues may be 'recruited' such that they influence the interpretation of an ambiguous rotating structure from motion (SFM) stimulus. Haijang et al. (2006) and Jain, Fuller, and Backus (2010) found that some cues (location, motion direction) were recruited as contextual cues that modulated SFM perception. However, they found that other cues (e.g. auditory cues and extrinsic visual cues) were not recruited to disambiguate the SFM stimulus (but see also Backus, Jain, & Fuller, 2011). We suggest that when a pair of cues has previously been unrelated in the environment, the visual system should represent this information; strong evidence that particular signals are unrelated allows the visual system to avoid learning new, spurious relationships. In contrast, learning will be faster when the visual system holds little information about whether or not two signals are correlated, or data that they are sometimes correlated. In this way, modification of cue-relationships may be better characterised as a continuum rather than by dichotomies such as parameter vs. structural learning (see Di Luca et al., 2010 for related discussion).

5. Conclusion

In summary, we show that colour can be learned and used as a cue to context: observers are able to selectively invoke different light priors in different contexts, allowing accurate recovery of shape from shading in their current environment. This use of context-dependent priors will assist the visual system as it moves between lighting environments, particularly when direct information about the prevailing illumination is unavailable or unreliable.

Acknowledgements

Thanks to Ben Backus for helpful scientific discussions. ISK was supported by an Economic and Social Research Council studentship.

References

Adams, W. J. (2007). A common light-prior for visual search, shape and reflectance. *Journal of Vision, 7*(11:11), 1–7.

Adams, W. J., Graf, E. W., & Ernst, M. O. (2004). Experience can change the 'light-from-above' prior. *Nature Neuroscience, 7*(10), 1057–1058.

Adams, W. J., Kerrigan, I. S., & Graf, E. W. (2010). Efficient visual recalibration from either visual or haptic feedback: The importance of being wrong. *Journal of Neuroscience, 30*(44), 14745–14749.

Adams, W. J., & Mamassian, P. (2004). Bayesian combination of ambiguous shape cues. *Journal of Vision, 4*(10), 921–929.

Backus, B. T. (2009). The mixture of bernoulli experts: A theory to quantify reliance on cues in dichotomous perceptual decisions. *Journal of Vision, 9*(1:), 1–19.

Backus, B. T. (2011). Recruitment of new visual cues for perceptual appearance. In J. Trommershäuser, K. Körding, & M. S. Landy (Eds.), *Sensory Cue Integration* (pp. 101–119). USA: Oxford University Press.

Backus, B. T., Jain, A., & Fuller, S. G. (2011). Cue recruitment for extrinsic signals after training with low-information stimuli. *Journal of Vision, 11*(11), 983 [Abstract].

Burge, J., Ernst, M. O., & Banks, M. S. (2008). The statistical determinants of adaptation rate in human reaching. *Journal of Vision, 8*(4:20), 1–19.

Di Luca, M., Ernst, M. O., & Backus, B. T. (2010). Learning to use an invisible visual signal for perception. *Current Biology, 20*, 1860–1863.

Ernst, M. O. (2007). Learning to integrate arbitrary signals from vision and touch. *Journal of Vision, 7*(5:7), 1–14.

Haijang, Q., Saunders, J. A., Stone, R. W., & Backus, B. T. (2006). Demonstration of cue recruitment: Change in visual appearance by means of Pavlovian conditioning. *Proceedings of the National Academy of Sciences, 103*(2), 483–488.

Hershberger, W. (1970). Attached-shadow orientation perceived as depth by chickens reared in an environment illuminated from below. *Journal of Comparative & Physiological Psychology, 73*(3), 407–411.

Honey, R. C., Bateson, P., & Horn, G. (1994). The role of stimulus comparison in perceptual learning: An investigation with the domestic chick. *Quarterly Journal of Experimental Psychology Section B, 47*(1), 83–103.

Jain, A., Fuller, S., & Backus, B. T. (2010). Absence of cue-recruitment for extrinsic signals: Sounds, spots, and swirling dots fail to influence perceived 3D rotation direction after training. *PLoS ONE, 5*(10), e13295.

Kleffner, D. A., & Ramachandran, V. S. (1992). On the perception of shape from shading. *Perception & Psychophysics, 52*(1), 18–36.

Knill, D. (2007). Learning Bayesian priors for depth perception. *Journal of Vision, 7*(8:13), 1–20.

Langer, M., & Bülthoff, H. H. (2001). A prior for global convexity in local shape from shading. *Perception, 30*(4), 403–410.

Michel, M. M., & Jacobs, R. A. (2007). Parameter learning but not structure learning: A Bayesian network model of constraints on early perceptual learning. *Journal of Vision, 7*(1:4), 1–18.

Mitchell, C., Nash, S., & Hall, G. (2008). The intermixed-blocked effect in human perceptual learning is not the consequence of trial spacing. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(1), 237–242.

Potetz, B., & Lee, T. S. (2003). Statistical correlations between two-dimensional images and three-dimensional structures in natural scenes. *Journal of the Optical Society of America A, 20*(7), 1292–1303.

Seydell, A., Knill, D. C., & Trommershäuser, J. (2010). Adapting internal statistical models for interpreting visual cues to depth. *Journal of Vision, 10*(4:1), 1–27.